

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

Gillis Maclean & Paul McKeown (2012): Comparing online quizzes and takehome assignments as formative assessments in a 100-level economics course, *New Zealand Economic Papers*, DOI:10.1080/00779954.2012.707530

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

Final submission

Abstract

Conventional take-home assignments and online quizzes are compared as formative assessments intended to engage students in learning. Using data from six semesters for each, we consider five characteristics: participation, timeliness and nature of feedback, fit within overall course assessment, and cost of delivery. Both assignments and quizzes generated high participation. Marked feedback took up to five weeks with assignments but was immediate with quizzes. In both cases, passing the formative assessment did not ensure a pass in the exam, but failing it indicated a lack of engagement and almost certain exam failure. The 10% course weighting for quizzes fitted better than the 30% for assignments. The assignments were costly to administer, but online quizzes had a marginal cost close to zero. As formative assessments, we find that overall online quizzes were as effective as take-home assignments and cost considerably less.

Gillis Maclean and Paul McKeown, Lincoln University

Key words: formative assessment, summative assessment, online assessment, economics education

JEL

A2 - Economics Education and Teaching of Economics

A22 - Undergraduate

Gillis Maclean

Economics Lecturer

Faculty of Commerce

P O Box 84

Lincoln University

Lincoln 7647

Christchurch, New Zealand

p + 64 3 4230219

e Gillis.Maclean@lincoln.ac.nz | w www.lincoln.ac.nz

1 Introduction

Assessing students' achievement of learning objectives is a core function of Tertiary Education Institutions (TEIs). As institutions, TEIs stand or fall on the quality of the assessments embodied in the degrees they award. Traditional forms of assessment are labour-intensive and TEIs face continual pressure to manage costs. Innovative solutions are appearing that reduce the time cost of feedback and assessment, such as interactive tutorials, online quizzes and better management of marks.

This paper compares the effectiveness of conventional hand-written, take-home assignments and online multiple-choice quizzes as partial assessments in a 100-level economics course at Lincoln University. Both were graded formative assessments, intended as frameworks in which students practise and learn by doing; students were allowed to use course materials and get help during the assessments. Grade-points were awarded as an incentive for engagement, with assignments worth 30% of the final grade and quizzes 10%. Over this time two invigilated tests and the final examination were the summative assessments, intended to measure students' achievement of learning objectives; together test and exam marks made up most of the final grade. As learning exercises, the take-home assignments and quizzes were not definitive assessments of achievement in the same way as the invigilated tests and exam. The formative assessments were complements to, rather than substitutes for, the summative assessments. Assessment is effective to the extent it works as intended. We consider five criteria for effective graded formative assessment: participation, timeliness and nature of feedback provided, fit within the overall course assessment, and cost of delivery (a consideration for any assessment).

Data from six semesters were analysed for both assignments and quizzes. Both assignments and quizzes successfully generated high levels of student participation, though participation in itself is a necessary but not sufficient condition for passing the course. Marked feedback took up to five weeks with assignments but was immediate with online quizzes; however, assignments provided more comprehensive feedback on written work, while quiz feedback on multiple-choice or calculations was limited to correct/incorrect. Passing either the assignments or quizzes did not ensure a pass in the exam, but failing either type of formative assessment indicated a lack of engagement and almost certain exam failure. The weighting of 10% for quizzes was found to fit better into the overall assessment programme than the 30% for assignments. While the assignments were costly to administer, online quizzes had a marginal cost close to zero. As formative assessments, we find that taken overall, online quizzes were as effective as take-home assignments and cost considerably less.

Section two describes the functions of the assessment used. Section three presents the data, Section four the results and Section five concludes.

2 Formative and summative assessment

There is increasing emphasis on assessment as TEIs respond to students' and employers' interest in the content of graduates' qualifications. Salemi, Siegfried, Sosin, Walstad and Watts (2001: 442) propose "improving the assessment of student learning in college economics courses" as one of their five initiatives for research in economic education. *Assessment* is an evaluation of any kind (Brookhart, 2001), though examiners focus most on *graded* assessment. There are typically two purposes for assessment. Formative assessment looks forward to

improve learning outcomes and summative assessment looks backward to assign final grades (Black & Wiliam, 2009; Brookhart, 2004; Miller, 2002; Nicol, 2007; Nicol & Macfarlane-Dick, 2006). Formative assessment is intended to inform learning and summative assessment to sum up learning. In the widest sense, formative assessment is any activity providing feedback which helps students learn. Most formative assessment takes place through informal evaluation without any marks being recorded; e.g. question and answer, tutorial participation or working through exercises with answers. Some formative assessment is formalised and graded as an incentive for study. Take-home assignments and online quizzes are typical, graded formative assessments: to encourage learning, such assessment is open-book, with little time pressure, and low stakes with few grade-points to avoid penalising students for mistakes. By contrast, the goal of summative assessment is to measure achievement of learning objectives at some point in the semester. This requires the assessment to discriminate across the range of outcomes from failing to excellent. The quintessential summative assessment is the final examination: closed-book and invigilated, time-limited and high stakes. Other typical, though less extensive, summative assessments include in-semester tests conducted under exam conditions.

While clear conceptually, the formative-summative distinction is not clear-cut when the assessment is marked (Brookhart, 2001 & 2004; Daly, Pachler, Mor & Mellar, 2010; Pachler, Mellar, Daly, Mor, Wiliam & Laurillard, 2009). The assignments and quizzes considered here are primarily intended as learning exercises (formative assessments), with marks as an incentive for participation; but the marks gained do count towards the final course mark, which is a summative assessment. This duality creates a tension, as discussed below (Section 4); marks awarded for participation in formative assessment may not be useful for summative purposes. Because of their differing contexts and functions, graded formative and summative assessments should be evaluated against different criteria. Their fit within the total assessment package in terms of relative weighting is a further consideration (for constructive alignment of assessment see Gillett & Hammond, 2009; Rust, 2002; Taylor, 2002; Treleaven & Voola, 2008). In particular, there will not necessarily be a close correlation between the marks from formative and summative assessments of the same topic. For the institution moderating a course like Econ 101, what counts are the outcomes – mean mark, grade distribution and pass rate – for the course as a whole only; the outcomes of individual pieces of assessment are not of interest in the big picture.

A recent conference comparing economic education across several countries shows that a range of assessment types are being used, many of which are intended as formative (Yamaoka, Walstad, Watts, Asano & Abe, 2010). For universities in New Zealand, Taiwan and the U.S.A., besides the final examination, common graded assessments include quizzes, homework/problem sets, written assignments, class participation, oral presentations, term papers, and mid-term examinations (Dalziel, Maclean & McKeown, 2010; Fann & Tsai, 2010; Schaur, Watts & Becker, 2008; Walstad & Watts, 2010). In Scotland, some more traditional universities have typically between 50% and 80% of assessment by examination, and other assessment by written assignments or group projects, with little use of multiple-choice tests (Fagan, 2010). Colander and McGoldrick (2009: 617) include implementing “summative and formative assessment of skill acquisition” in their proposals for improving pedagogical practices. Smith, Clements and Olson (2010) and Weldy and Turnipseed (2010) consider the impact of professional accreditation on assessment in business education.

The literature on formative assessment requires attention to detail. Some assessment described as formative is arguably instead small scale summative assessment, and words like *exam* have different meanings in context. Student behaviour and outcomes depend critically on the

assessment set-up, including how many grade-points are at stake, how marks are awarded, how many attempts are allowed, and whether answers are provided for incorrect responses. Nonetheless, other things being equal, the medium (on-paper or online) of the graded formative assessment (assignment, homework, quizzes) appears to make little difference to summative outcomes.

Few studies on graded formative assessment in economics directly compare the online with the traditional on-paper medium. Kennelly, Considine and Flannery (2011), and Lee, Courtney and Balassi (2010) found that changing the formative assessment medium (from on-paper to online) made no significant difference to (summative) exam marks; Palocsay and Stevens (2008) found the same in business statistics. Without reference to on-paper formative assessment, Galizzi (2010) found no significant effect, and Stephenson (2001) found a “modest” increase, in exam marks after the introduction of online formative assessment. Against this there are less convincing claims that using online homework systems resulted in increased exam or course marks in economics, accounting and finance (Collins, Deck & McCrickard, 2008; Nguyen & Trimarchi, 2010; Pace, 2010); the first two studies did not control for other factors like changes in the exam or class composition, and the third provided no evidence at all.

3 Assessment in Econ 101/110, 2001-2011: setup and data

Both authors have taught Lincoln University’s entry-level economics since 2001. This one-semester course covers both microeconomics and macroeconomics as part of the compulsory core for several degree programmes. Like most courses, this evolved over 2001-2011, with a syllabus revision, and change in title from Econ 101 (Principles of Economics) to Econ 110 (Introduction to Applied Economics). We consider only the periods 2001-03 with assignments and 2008-11 with quizzes. The course design and teaching culture remained substantially the same, with consistent student outcomes; the overall course mean mark and pass rate were almost identical over the two relevant periods (see Table 2).

The primary purpose of the graded formative assessment was to get students engaged in learning throughout the semester, as opposed to cramming just before their tests and exam, and to provide reasonably continuous feedback on their progress to both student and teacher. Most students study four courses which compete for their time. In our experience, students tend to allocate their attention across their courses according to the relative grade-points at stake in the very near future. Tasks that don’t carry grade-points are likely to be neglected. The grade-points awarded for formative assessment in Econ 101/110 were thus primarily incentives for participation and, in retrospect, proved to have limited usefulness for summative purposes.

As exercises, the assignments and quizzes had three things in common. First, they were open-book and not invigilated, the purpose being to get students using learning resources and seeking help. Without invigilation, graded assessment is open to plagiarism and cheating (e.g. Harmon & Lambrinos, 2008; Judge, 2008). We found very little evidence of cheating in the take-home assignments. However, copying was hard to detect given the combination of multiple markers and a substantial proportion of numerical or graphical questions with only one possible correct answer. We did not invigilate the quizzes. Students must log in to submit a quiz, though it is possible another person could answer. But the maximum grade-points earned from any one quiz was 0.83%, which we judge to be little incentive to cheat. Second, there was no race against the clock as in tests and exams. Third, they were low stakes (5% and 0.83% per item respectively)

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

so students were not harshly penalized for mistakes, though this also meant most students would expend less effort than in the tests and exam with more grade-points.

In stark contrast, the summative assessment was closed-book, time-limited and high stakes. Given the polarity in conditions, even if an identical question is asked, the formative and summative assessments evaluate the outcome of quite different activities. For this reason, marks from the assignments or quizzes were not a reliable predictor of marks from the summative assessment, and the exam in particular (a consequence evident to us in hindsight). If a student can correctly answer with help and ample time, it does not follow they can do the same under exam conditions. Rather, passing the formative assessment is close to being a necessary, but not sufficient condition for passing the summative. As an analogy, consider a task where the objective is to run 100 metres, with a pass awarded for completion within a target time, set at one minute for practice sessions, and twenty seconds for the race. A pass in practice does not ensure passing the race, but a fail in practice means almost certainly failing the race.

Over 2001-2003, coursework consisted of five take-home written assignments (together worth 30%) and two invigilated multiple-choice-only tests (together worth 20%). The final exam was 50% of the final grade. The assignment questions followed the format of the final exam. The exercise was designed to allow students to learn by doing before being marked. Students got feedback from interacting with learning resources throughout the exercise. They were expected to work from the textbook and lecture notes; similar questions were covered in tutorials; they were encouraged to seek help, and even get feedback on their draft before submitting their final copy. That students seek out assistance is a sign that the formative assessment works as intended. After two weeks' lectures, students started the assignment with two weeks before submission. Since this task could be completed in several hours, there was little time pressure. Given the help available and the opportunity to correct mistakes *before* assignments were marked, any student applying effort could, and should, pass. Marked scripts were returned one week later. The complete process took five weeks from the first corresponding lecture. The same topics were retested in the final exam, though the room for improvement from the marked assignment was limited by the high assignment average of 73%.

The assignments were stopped in 2004 because they were costly to run and did not function as well as expected (see Tables 2 and 3). With no obvious replacement for assignments, there were no graded formative assessments until 2008 when online assessment became feasible. Online quizzes in the Moodle learning management system were trialled in 2008 semester 1 without grade-points. In 2008 semester 2, graded formative assessment was re-introduced through online quizzes together worth 10% of the final grade. There were 12 quizzes, each covering one week's material. Each quiz had 10 questions selected randomly from a bank of around 30 questions; most were multiple-choice, though some quizzes included calculations. The learning context was the same as for assignments, as the assessment was open-book, not a race against the clock and low-stakes. With each quiz open for between two and four weeks, and the mean total time spent on each quiz 91 minutes, there was little time pressure. (No reliable data are available for time taken writing assignments; the capacity of online learning management systems to record data on student behaviour is a significant innovation.) While any quiz was open, students could make unlimited attempts, with only the best mark recorded as an incentive to keep practising. Each submitted quiz got immediate feedback showing whether each response was correct or incorrect. Correct answers were not supplied for incorrect responses, as this would lead to correct answers being copied for further attempts. Students wanting to increase their marks had to work out their mistakes or ask for help (or resort to brute force and submit random responses). Under these conditions, any student applying effort could, and

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

should, pass. The same topics were retested in the multiple-choice sections of the tests and exam.

From 2008, students had extensive weekly exercises with answer guides in the course handbook. This actually gave students more opportunities for formative assessment of written work than the assignments, albeit without grade-points. With the change to quizzes the assessment package was recalibrated for better balance, with the tests more challenging than before and the exam less so. The two tests were expanded to include written answers, and their weighting increased to 40% of the final grade (from 20% for the previous multiple-choice-only tests). For the summative assessment of written answers the invigilated tests had two important advantages over take-home assignments; the assessment was secure, and two tests cost less to write and mark than five assignments.

To evaluate the effectiveness of the formative assessment we took data from Econ 101/110 in the two 12-week semesters each year, with six semesters for each assessment over the calendar years 2001-03 and from 2008 semester 2 to 2011 semester 1. We consider only students who sat the exam; students with an aegrotat for any relevant assessment are excluded because of impaired performance. Repeating students are counted in all semesters that they appear. No scaling was used and all marks were converted to percentages. A mark (after rounding) of 50% or greater is a pass. “Assignments” refers to the total mark for all assignments, “quizzes” refers to the total mark for all quizzes, and “exam” refers to the mark for the final exam. For assignments, the population is all students ($n = 1802$) over six semesters who completed at least one assignment and sat the final exam. For quizzes, the population is all students ($n = 1204$) over six semesters who completed at least one quiz and sat the final exam.

For simplicity we compare assignments and quizzes against the final exam without considering the tests. The final exam is the most comprehensive summative assessment and the final arbiter of grades in Econ 101/110 with 50% of the total marks in both periods. The exam format, scope and time allowed were kept constant. From 2008 however, to get the desired overall course outcomes, the exam was recalibrated. By design the exam mean and pass rate differ significantly for the two periods, while the course outcomes are consistent (Table 2). This means that, unlike the studies cited above (end of Section 2), it is not possible to test whether students completing assignments (2001-3) scored more or less in their respective exams than students completing quizzes (2008-11).

4 Formative assessment in Econ 101/110, 2001-2011: results

One purpose of the graded formative assessment was to keep students engaged throughout the semester by providing incentives to practise. Both assignments and quizzes were very effective in generating participation, as shown in Table 1 (note: in Table 1 only, for participation the populations include students with no attempts at assignments or quizzes.) Over three years, of all students sitting the final exam only 0.1% did not attempt any assignments, and 93% submitted at least four of the five assignments. The Christchurch earthquake (February 2011) meant the first semester was shortened, and 2011 semester 1 is omitted from the quiz results as incomparable for this reason. Over the five remaining semesters, of all students sitting the final exam, only 1% did not attempt any quizzes and 93% submitted at least seven of twelve quizzes. There is little to distinguish assignments from quizzes in terms of participation rates. It is interesting that dropping the course weighting from 30% (assignments) to 10% (quizzes) did

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

not result in an appreciable drop in participation rates despite the incentive being cut to a third. This suggests the 30% weighting was unnecessarily generous and that 10% is sufficient as an incentive for participation.

Table 1: participation rates for assignments and quizzes

| Assignments 2001-2003 | | Quizzes 2008-2010 | |
|----------------------------|------------|-----------------------------|------------|
| number submitted (of 5) | % students | number submitted (of 12) | % students |
| 0-2 | 2.8% | 0-3 | 3.1% |
| 3 | 4.0% | 4-6 | 3.8% |
| 4 | 11.6% | 7-9 | 6.6% |
| 5 | 81.6% | 10-12 | 86.5% |
| | 100% | | 100% |

We did not observe any significant difference between the assignments and quizzes in the number of students who took the opportunity to seek advice. In student surveys 2008-11 both the course handbook and quizzes were rated highly as learning resources. We guess that five assignments took between ten and twenty hours per semester, and we know that twelve quizzes took on average eighteen hours. This suggests the quizzes were at least as effective as the assignments for the average time of engagement over the semester, and that as a strategy, it may be better to provide smaller pieces more often than larger pieces less often.

Another purpose of the graded formative assessment was to provide reasonably continuous feedback on progress to both student and teacher. Because of the differences in intent and context, the marks from the formative assessments are not closely correlated with the marks from the summative assessments and provide less discrimination across the range of outcomes from failing to excellent. Formative assessment complements, rather than substitutes for, summative assessments. For the examiner there is a trade off: attaching more weight to formative assessment may encourage more practice but the reduction in summative assessment reduces the separation across the range of grades. Hence the formative assessment weighting has implications for the overall assessment regime.

The scatter diagrams in Figures 1 and 2 show marks for the exam against the assessment in question; each point represents one student. The solid lines are the respective mean marks and the $y = x$ line. If the two pieces of assessment were of equal difficulty, marks would be equally distributed above and below the $y = x$ line. In both cases, the bulk of points are below the $y = x$ line, showing that assignment or quiz marks are much easier to get than exam marks. While assignment/exam marks cluster around their means, quiz/exam marks are skewed with 26% of all students scoring the maximum 100 for quizzes.

Fig.1: assignments vs final exam (n = 1802)

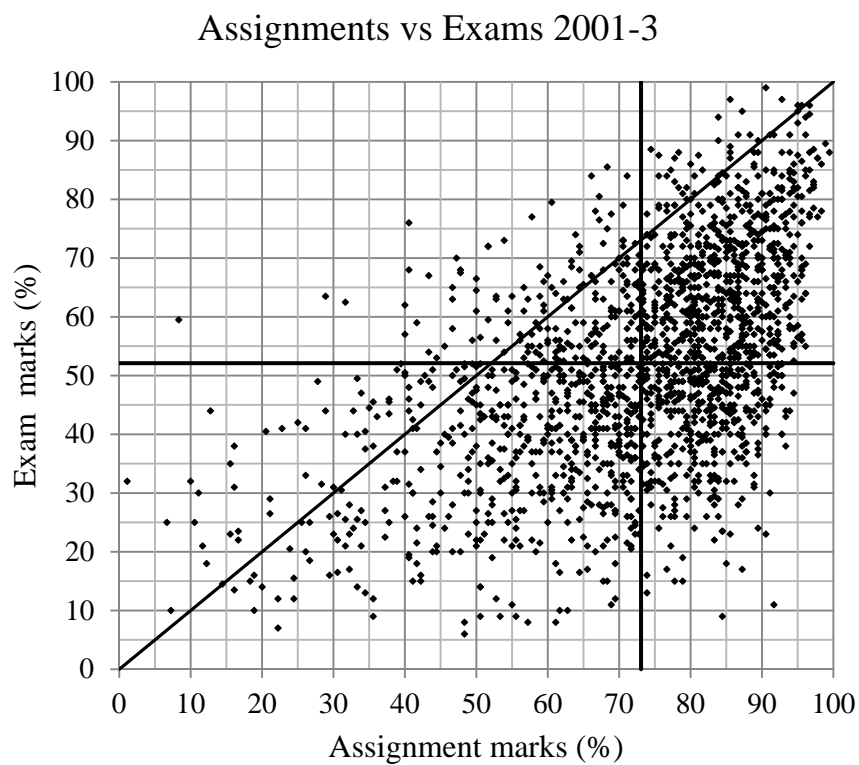


Fig.2: quizzes vs final exam (n = 1204)

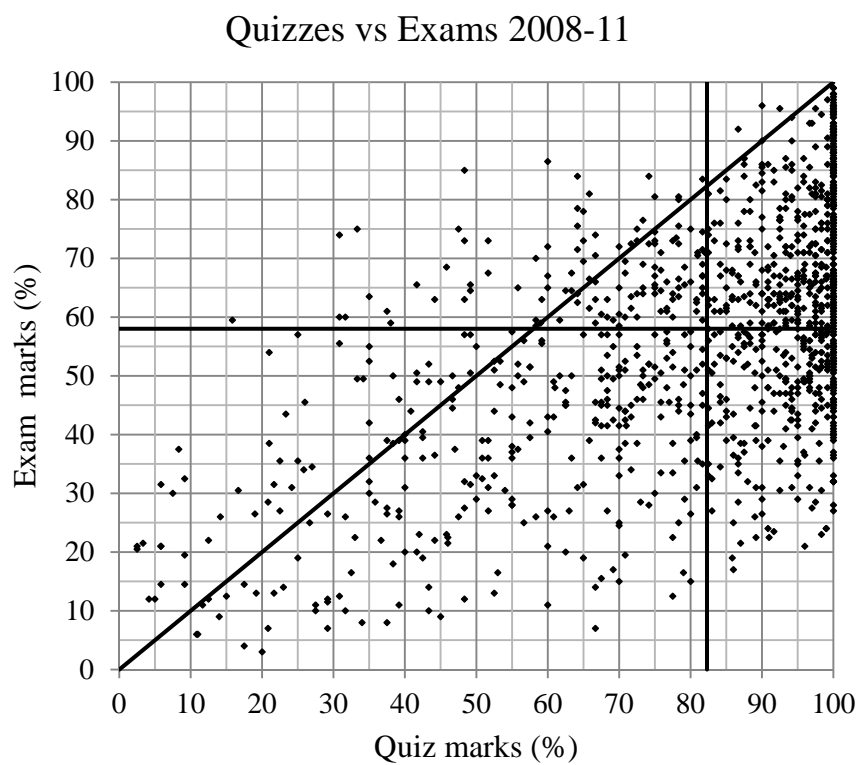


Table 2 shows consistent mean total marks and course pass rates over both periods. With hindsight, the 2001-03 assessment regime did not work as intended primarily because assignments were overweighted at 30% of the course mark. This “grade-point inflation” of assignments made it difficult to meet an acceptable final grade distribution and pass rate. As the mean assignment mark was 22/30 (73%), and the mean test mark 12/20 (62%), an average student needed only 16/50 (32%) from the exam to reach 50% overall and pass the course. A defensible course pass rate then required setting a more challenging exam and imposing a minimum exam mark of 40% for a clear pass. The assignments were stopped in 2004 and without them, the minimum exam mark was no longer required. We note that the problem lay in the overweighting of the assignment marks rather than in the assignment itself as an exercise.

Table 2: graded assessment summary Econ 101/110

| | 2001-2003 | | | 2008-2011 | | |
|------------|-----------|------|-----------|-----------|------|-----------|
| | Weight | Mean | Pass rate | Weight | Mean | Pass rate |
| Formative | 30% | 73 | 90% | 10% | 82 | 89% |
| Tests | 20% | 62 | 77% | 40% | 58 | 65% |
| Exam | 50% | 52 | 56% | 50% | 58 | 68% |
| Total mark | 100% | 61 | 74% | 100% | 61 | 75% |

Comparison between assignments and quizzes is complicated by offsetting changes to other assessment, as the overall assessment regime was recalibrated, with the goal of maintaining consistent course outcomes (mean mark and pass rates). As tests were made more comprehensive by changing from shorter, multiple-choice-only tests (2001-03) to longer tests including written answers (2008-11), the tests mean and pass rate were reduced. On the other hand, when the weight for the formative assessment was reduced from 30% to 10%, the exam was no longer required to be so severe, and the exam mean and pass rate were increased. We think the 2008-11 regime gives a better balance between tests and the exam.

The pass rates for assignments and quizzes were 90% and 89% respectively, much higher than the respective exam pass rates of 56% and 68%. Both assignment and quiz means (73 and 82) were much higher than the respective exam means (52 and 58). The consequences of these disparities are described in Table 3.

Table 3: assignments & quizzes vs final exam

| | Assignments 2001-2003 | Quizzes 2008-2011 |
|---|--------------------------|----------------------|
| Conditional probability of ... | | |
| passing exam given pass in assignment/quizzes | 60% | 74% |
| failing exam given fail in assignment/quizzes | 82% | 77% |
| % of population who ... | | |
| scored more in exam than assignment/quizzes | 10% | 9% |
| passed exam after failing assignment/quizzes | 2% | 3% |

Table 3 shows that of those students who had already passed the assignments or quizzes, only 60% and 74% respectively passed the subsequent exam. This difference in the probabilities of exam pass/fails can be largely explained by the higher exam pass rate 2008-11 when quizzes were used. A pass in the assignments or quizzes thus contains little information about the likely performance in the subsequent exam, and can be taken more as a signal that a student has continued to have some engagement in the course rather than dropping out. More significantly,

only 10% and 9% of students scored more in the exam than in the assignments and quizzes respectively, and only 2% and 3% of the population passed the exam after failing the assignments and quizzes respectively. Thus passing the formative assessment was (almost) necessary but not sufficient to ensure an exam pass, and failure in the formative assessment indicated a lack of engagement that almost certainly resulted in exam failure. We agree with O’Byrne and Thompson (2005: 84) that “a lack of good correlation between in-semester assessment and examination results is typical in our experience.”

The assignments consisted of constructed response questions, which made up 71% of the exam. The quizzes consisted of multiple-choice questions, which made up 29% of the exam. Based on their format we assumed that assignments provided more comprehensive and better quality feedback in preparation for the summative assessment. However, the results do not support this supposition. There is no evidence that assignments functioned better than quizzes with respect to the subsequent exam. This is an interesting result. In Econ 101/110, whether or not students were engaged in some kind of formative assessment process turns out to be more important than whether that assessment was written or multiple-choice. This is in line with the conclusions of Kennelly, Considine and Flannery, (2011), Lee, Courtney, and Balassi, (2010) and Palocsay and Stevens, (2008).

The first critical difference is the timeliness of the feedback for both student and teacher. With assignments students practised material two or three weeks after the corresponding lectures, and then scripts took a week to mark. The process took up to five weeks, which could leave little time to overcome any deficiencies revealed. With quizzes students practised one week after lectures, and marking took seconds. Another important innovation of the online quiz is the ability to monitor students in real time and provide intervention much earlier than previously possible. Further, digital media such as online quizzes create activity logs that can be used to measure student learning behaviours that would otherwise be unobservable.

The second critical difference is the delivery cost. Written assignments are expensive. Each semester requires writing a new set of assignments and marking is time consuming with costs proportionate to class size. Further, writing new assignments each semester and using multiple markers for timely return of scripts, creates problems for consistency in assessment across semesters. In contrast, the cost of importing a test bank into the learning management system is minimal, and then the marginal cost of online quizzes is close to zero regardless of class size. Consistency is given as random ordering allows the same quiz banks to be reused every semester, and marking is automated.

5 Conclusion

We compare two formative assessments, the online multiple-choice quiz and the traditional take-home assignment, with six semesters of data for each. Both were open-book, not races against the clock, and carried few grade-points per item. As learning exercises, they were not definitive assessments of student achievement in the same way as the invigilated tests and exam, which were closed-book, time-limited and high stakes. The formative assessments were complements to, rather than substitutes for, the summative assessments. Five criteria for effective graded formative assessment were considered.

Participation can be measured by the percentage of the class engaged, and the average time students spent engaged over the semester. Both measures suggest the quizzes were as effective

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

as the assignments in getting engagement, despite having only one third of the assignments' grade-points as an incentive. This suggests the 30% weighting was unnecessarily generous and that 10% is sufficient as an incentive for participation. It may be better to provide smaller pieces more often than larger pieces less often. The first critical difference is the timeliness of the feedback. Assignments lagged corresponding lectures by two or three weeks and scripts took a week to mark. Quizzes lagged lectures by one week and were marked in seconds. Another important innovation is that the rapid response of online quizzes allows the possibility of finding learning problems and intervening much earlier than was possible with manually marked assignments. Both assignments and quizzes engaged students with learning resources as they practised. Given their difference in context, formative assessment marks were not correlated with exam marks. Passing either the assignments or quizzes did not ensure a pass in the exam, but failing either the assignments or quizzes indicated a lack of engagement and almost certain exam failure. There was no evidence that on-paper assignments functioned better than online quizzes with respect to the subsequent exam, a result in line with most studies cited above (Section 2). Further, whether or not students in Econ 101/110 were engaged in some kind of formative assessment process turns out to be more important than whether that assessment was written or multiple-choice.

Independently of the above considerations, as to the weighting of formative assessment, 10% is a better fit within the course than 30%. This is not the critical point in comparing assignments with quizzes. There is nothing to suggest that results would have been different if the weighting for assignments and quizzes had been reversed. The second critical difference is the delivery cost. Assignments were costly to write and mark while online quizzes had a marginal cost close to zero. As formative assessments, we find that overall, online quizzes were as effective as take-home assignments and cost considerably less.

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

Black, P. and Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31.

Brookhart, S.M. (2001). Successful students' formative and summative uses of assessment. *Assessment in Education: Principles, Policy & Practice*, 8, 153-169.

Brookhart, S.M. (2004). Assessment theory for college classrooms. *New Directions for Teaching and Learning*, 100, 5-14.

Colander, D. & McGoldrick, K. (2009). The economics major as part of a liberal education: the Teagle report. *American Economic Review*, 99, 611-18.

Collins, D., Deck, A. & McCrickard, M. (2008). Computer aided instruction: a study of student evaluations and academic performance. *Journal of College Teaching & Learning*, 5, 49- 57.

Daly, C., Pachler, N., Mor, Y. & Mellar, H. (2010). Exploring formative eassessment: using case stories and design patterns. *Assessment & Evaluation in Higher Education*, 35, 619-636

Dalziel, P., Maclean, G. & McKeown, P. (2010). Teaching economics in New Zealand: recent trends and issues. In Yamaoka, Walstad, Watts, Asano & Abe (Eds.) (2010), *Comparative studies on economic education in Asia-Pacific region* (pp.185-207). Tokyo: The Institute of Asia-Pacific Studies, Waseda University.

Fagan, C. (2010). Economics education within the education system in Scotland. In Yamaoka, Walstad, Watts, Asano & Abe (Eds) (2010), *Comparative studies on economic education in Asia-Pacific region* (pp.97-117). Tokyo: The Institute of Asia-Pacific Studies, Waseda University.

Fann, G.J. & Tsai, C.L. (2010). Economics education in Taiwan. In Yamaoka, Walstad, Watts, Asano & Abe (Eds.) (2010), *Comparative studies on economic education in Asia-Pacific region* (pp.119-155). Tokyo: The Institute of Asia-Pacific Studies, Waseda University.

Galizzi, M. (2010). An assessment of the impact of online quizzes and textbook resources on students' learning. *International Review of Economics Education*, 9, 31-43.

Gillett, A. and Hammond, A. (2009). Mapping the maze of assessment: An investigation into practice. *Active Learning in Higher Education*, 10, 120–137.

Harmon, O.R. & Lambrinos, J. (2008). Are online exams an invitation to cheat? *Journal of Economic Education*, 39, 116-125.

Judge, G. (2008). Plagiarism: bringing economics and education together (with a little help from IT). *Computers in Higher Education Economics Review*, 20, 21-26.

Kennelly, B., Considine, J. and Flannery, D. (2011). Online assignments in economics: a test of their effectiveness. *The Journal of Economic Education*, 42, 136-146.

Lee, W., Courtney, R.H., and Balassi, S.J. (2010). Do online homework tools improve student results in principles of microeconomics courses? *American Economic Review*, 100, 283-86.

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

Miller, N. (2002). Alternative forms of formative and summative assessment. *The handbook for economics lecturers*. The Economics Network.

<http://www.economicsnetwork.ac.uk/handbook/assessment/welcome>

Nguyen, T. & Trimarchi, A. (2010). Active learning in introductory economics: do MyEconLab and Aplia make any difference? *International Journal for the Scholarship of Teaching and Learning*, 4, 1-18.

Nicol, D. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31, 199–218.

Nicol, D. (2007). Principles of good assessment and feedback: theory and practice. From the REAP International Online Conference on Assessment Design for Learner Responsibility, 29th-31st May, 2007. Available at <http://ewds.strath.ac.uk/REAP07>

O’Byrne, J. & Thompson, R. (2005). The tutorial benefits of on-line assignments: mastering physics in first year physics at the University of Sydney. In *Proceedings of Symposium 'Blended Learning in Science Teaching and Learning'*, UniServe Science, Sydney 2005, 81-86.

Pace, R.D. 2010. Using Aplia in finance classes: student perspectives and performance. *American Journal of Business Education*, 3, 63- 68.

Pachler, N., Mellar, H., Daly, C., Mor, Y., Wiliam, D. & Laurillard, D. (2009). *Scoping a vision for formative e-assessment: A project report for JISC*. London: Institute of Education. www.jisc.ac.uk/media/documents/projects/scopingfinalreport.pdf

Palocsay, S.W. & Stevens, S.P, (2008). A study of the effectiveness of web-based homework in teaching undergraduate business statistics. *Decision Sciences Journal of Innovative Education*, 6: 213-232.

Rust, C. (2002). The impact of assessment on student learning: How can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices? *Active Learning in Higher Education*, 3; 145–158.

Salemi, M.K., Siegfried, J.J., Sosin, K., Walstad, W.B. & Watts, M. (2001). Research in economic education: five new initiatives. *American Economic Review*, 91, 440-45.

Schaur, G., Watts, M. & Becker, W.E. (2008). Assessment practices and trends in undergraduate economics courses. *American Economic Review*, 98, 552–556.

Smith, L.J., Clements, C. & Olson, J.E. (2010). Local versus standardized content assessment: some management implications, or why bother? *Journal of Education for Business*, 85, 249-257.

Stephenson, K. (2001). Assessments of the educational value of web-based instructional tools for introductory agricultural economics. *Review of Agricultural Economics*, 23, 492–509.

Comparing online quizzes and take-home assignments as formative assessments in a 100-level economics course

Taylor, R. (2002). Designing Undergraduate Degree Programmes. *The handbook for economics lecturers*. The Economics Network.

http://www.economicsnetwork.ac.uk/handbook/printable/programmes_v5.pdf

Treleaven, L. and Voola, R. (2008). Integrating the development of graduate attributes through constructive alignment. *Journal of Marketing Education*, 30, 160-173. DOI: 10.1177/0273475308319352

Walstad, W.B. & Watts, M.W. (2010). Undergraduate economics in the United States. In Yamaoka, Walstad, Watts, Asano & Abe (Eds.), (2010), *Comparative studies on economic education in Asia-Pacific region* (pp.157-183). Tokyo: The Institute of Asia-Pacific Studies, Waseda University.

Weldy, T.G. & Turnipseed, D.L. (2010). Assessing and improving learning in business schools: direct and indirect measures of learning. *Journal of Education for Business*, 85, 268-273.

Yamaoka, M., Walstad, W.B., Watts, M.W., Asano, T. & Abe, S. (Eds.). (2010). *Comparative studies on economic education in Asia-Pacific region*. Tokyo: The Institute of Asia-Pacific Studies, Waseda University.